A classroom-based study of Chinese as a Foreign Language pronunciation targeting syllables with final '-i'

By

Andrew S. Ring

Submitted to the graduate degree program in Curriculum and Teaching and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Master of Arts.

_____

Chairperson Dr. Manuela Gonzalez-Bueno

_____

Dr. Marc Mahlios

_____

Dr. Suzanne Rice

Date Defended: March 27, 2013

The Thesis Committee for Andrew S. Ring
certifies that this is the approved version of the following thesis:

A classroom-based study of Chinese as a Foreign Language pronunciation targeting syllables with final '-i'

_____

Chairperson Dr. Manuela Gonzalez-Bueno

Date approved: March 27, 2013

As a less commonly taught language in the U.S., relatively little research has been done on the acquisition of segmental sounds in Chinese as a Foreign Language (CFL), particularly at the secondary level. This study targeted a unique set of status-bearing syllables in Chinese described in the literature as potentially difficult for English-L1 learners by measuring the perceived and productive levels of difficulty for these sounds in 30 secondary-level CFL learners. An error analysis of learner performance on an oral reading assessment revealed two syllables that appeared to present difficulties for these learners. The range of phonological processes that occurred for these two particular syllables are presented along with the results indicating where perceived level of difficulty for these syllables converged or diverged with productive difficulties. In light of the results, a discussion of practical implications for teaching and learning is also provided.

INTRODUCTION

With China's increasing prominence in the global economy, the field of Chinese foreign language (CFL) teaching in the U.S. has enjoyed an unprecedented opportunity for growth especially in the last decade. (Ke, 2010) For the first time, Chinese courses are being made more widely available to students prior to entering college. A summary of recent research on the effective of foreign language pronunciation instruction suggests that explicit instruction seems to benefit learners, although depending on personal factors, students may tend to perceive it as more useful than teachers. Currently Communicative Language Teaching (CLT) approaches tend to advocate for comprehensibility and fluency as learning goals rather than native-like accuracy. Naturally, this perspective favors a more global focus over a purely segmental one. Notwithstanding, this study is primarily an attempt to account for the difficulties one group of high school English-speakers encountered with a rather unique status-bearing group of syllables at different stages of learning Chinese as a foreign language (CFL). These syllables are typically categorized by the articulatory locus of their initial consonant sounds as the dentals (/tsɿ/, /ts'ɿ/, /sɿ/), palatals (/tɕi/, tɕ'i /, /ɕi/), and retroflexes (/tʂl̩/, / tʂ'l̩/, /ʂl̩/, /ʐl̩/) (Chao, 1968; Dow, 1972; Duanmu, 2007; Hartman, 1944; Norman, 1989; Sun, 2006; Thompson and Li, 1987; Xing, 2006). That their segmental sounds, i.e. consonants and vowels, pose difficulties for CFL learners has been substantiated not only in the literature (Dow, 1972; Norman, 1989; Xing, 2006; Shi and Wen, 2009) but also by over a decade of personal and professional experience as a second language learner, instructor, and researcher in China, Taiwan and the United States. Through personal inquiry related to the apical vowels occurring in the dentals and retroflexes, some native Chinese language teachers from Mainland China,

where Hanyu Pinyin (Pinyin) is used as a supplementary spelling system to transcribe the sounds of Chinese, have described these vowels as being dependent on their consonants for realization. Some teachers from Taiwan, where a system roughly equivalent to Pinyin but relying on a stylistically Chinese alphabet, called Zhuyinfuhao is used, however, have described these syllables as not containing vowels. See Table 1 below for a conversion chart of the target sounds.

Table 1. Conversion chart of ten target sounds

	Dental			Retroflex				Palatal		
IPA	tsɿ	tsʾɿ	sɿ	tʂɻ	tʂʾɻ	ʂɻ	ʐɻ	tɕi	tɕʾi	ɕi
Pinyin	zi	ci	si	zhi	chi	shi	ri	ji	qi	xi
Zhuyinfuhao	ㄗ	ㄘ	ㄙ	ㄓ	ㄔ	ㄕ	ㄖ	ㄐ ㄧ	ㄑ ㄧ	ㄒ ㄧ

In learning Chinese, a learner is not only acquiring its target set of individual sounds but is also acquiring its sound system that determines the relationships among sounds (Hua and Dodd, 2000). As a CFL instructor seeking ways to make CFL more accessible, the topic's immediate relevance for classroom teachers and learners gave it practical appeal.

LITERATURE REVIEW

Chinese Phonology from Multiple Theoretical Perspectives

Standard Chinese (SC), also referred to as, among other things, Mandarin Chinese, Mandarin, and Putonghua is said to have a relatively simple syllabic structure due to the fact that Chinese does not permit consonant clustering in either initial or coda positions (Sun, 2006; Xing, 2006; Duanmu, 2007). Chinese initial consonants are considerably more constrained than those in English. Many of them combine only with a limited set of the vowels that are possible in Chinese. For example, whereas the dentals and retroflexes never combine with the high front vowels /i/ and /y/, the palatals never combine with

anything but /i/ and /y/. Although coda consonants are permissible, they are restricted to three: two nasals, one dental, /_n/, the other velar, /_ŋ/, and a retroflex, /_ɹ/.

One complicating feature of the Chinese syllable is tone. Tone is represented orthographically in Hanyu Pinyin, the Roman alphabetic spelling system most widely adopted as an aid for literacy instruction, hereafter referred to simply as Pinyin, by placing a diacritic mark above the main vowel, or the strong vowel as opposed to the medial vowel or gliding element(s) in a diphthong or triphthong. Although the focus of this study is on consonants and final vowels, which have traditionally been viewed as segments, tone is treated for convenience as a supersegmental feature. It is worth noting, however, that in Duanmu's (2007) analysis of Chinese tone, he argues that for consistency it is better to treat tone as a segmental feature, potentially consisting of laryngeal features, marked as [+/-thick] and [+/-tense]. Tones, therefore, in his view cannot be considered phonemic just as other segmental features are not considered phonemic, although they are a rich source for minimal contrasts of monosyllabic words, such as 'zhi-1' (知／知 ：to know), 'zhi-2' (直／直：straight), 'zhi-3' (只／只：only), 'zhi-4' (治／治：to govern or cure).

There are also disagreements over phonemic inventories in Chinese. One main source of disagreement is related to the sounds targeted in this study, particularly with whether the dental ('zi' /tsɿ/, 'ci' /ts'ɿ/, 'si' /sɿ/) and retroflex groups ('zhi' /tʂʅ/, 'chi' /tʂ'ʅ/, 'shi' /ʂʅ/, 'ri' /ʐʅ/) contain their own unique vowels or whether they should be considered vowel-less, syllabic consonants. Traditional interpretations consider dental and retroflex vowels to be allophonic variations of /i/ and transcribe them phonetically with apical vowels [ɿ] and [ʅ] or [z] and [ʐ], respectively (Dow, 1972; Thompson and Li, 1987;

Norman, 1988; Hua and Dodd, 2000). Duanmu (2007) argues to treat them as syllabic consonants primarily because the dental and retroflex syllables are unique in that they do not seem to rhyme. While this interpretation has the advantage of a reduced inventory of phonemes, it contradicts the notion of the vowel being the obligatory element in the Chinese syllable and also forces a reinterpretation of where tone occurs in the syllable. Howie (1976), in contrast, provides phonetic evidence of representing these as close, but distinct high, central vowels.

In sum, generally it is argued that the apical sounds occurring in the dental and retroflex syllables are allophonic variants of /i/, rather than phonemic, because they clearly occur in complementary distribution with the palatal syllables (ji /tɕi/, qi /tɕ'i/, xi /ɕi/). There also seems to be consensus with regard to articulation of these sounds. Following Norman and others, the traditional notations for the apical vowels co-occurring with dental and retroflex consonants [ɿ] and [ʅ], respectively, have been adopted for this study. From a pedagogical perspective, because popular Chinese language textbooks used in the United States adopt Pinyin rather than alternative orthographic systems and conventionally, Pinyin represents these syllables as containing the vowels, '-i', which bear tones, for the purpose of this study, the apical vowels are treated as vowels.

One sound for which there has been slight disagreement over manner of articulation is the retroflex, 'ri' /ʐʅ/. Some scholars, such as Norman (1988), have defined it as an approximant [ɻ], while most, such as Sun (2006) and Duanmu (2007) consider it a voiced fricative [ʐ]. Norman (1988) claims, "(T)he Chinese 'r' is pronounced with less friction than the comparable English fricative (/ʒ/), and acoustically sounds much closer to the usual American pronunciation of 'r' (/ɻ /). Moreover, to consider it the voiced counterpart of 'sh'

(/ʃ/) would be tantamount to recognizing voicing as a distinctive feature in the phonological system of Chinese, a distinction which is otherwise unneeded" (p. 140). In response, Duanmu (2007) has argued that /ʐ/ is a more faithful representation than /ɹ/, because the relationship between /ʂ/ and /ʐ/ is similar to that between /s/ and /z/, contrastive voicing. Dow (1972) and Sun (2006) have also distinguished /ʐ/ from the rhotacized vowel /_ɹ/, which would otherwise result in the rhotacized vowel being misrepresented as phonetically equivalent with the voiced, fricative, retroflex consonant. Dow in his articulatory phonetic manual described several differences between the English /ɹ/ and the Chinese /ʐ/, including length, friction, tenseness, location, and lip movement, while admitting more similarity between the two sounds when preceding rounded vowels, e.g. /u/. Duanmu (2007) and Chen (1999) both describe regional variation among native speakers of the dental and retroflex series, to the latter of which /ʐ/ belongs, suggesting the inadvisability of rigid standards with regard to this sound. For pedagogical purposes, /ʐ/ was adopted for this study as more natural and target-like with respect to SC.

Conceptual Framework

A number of second language acquisition (SLA) theories have significantly influenced the present research on and understanding of the acquisition of the target sounds. An account of those theories referred to in previous research and used in interpreting the present research is provided here.

Early in the development of the field of SLA, the Constrastive Analysis Hypothesis was developed within the Behaviorist Framework. The assumptions were that language learning is habit formation and that the differences between the learner's native language

(NL) and the target language (TL) would be the source of errors in learner output. Two versions of CAH have been considered: the strong, predictive or apriori version and the weak, explanatory or a posteriori version. The former version of CAH, as its name suggests, predicts that the greater the differences between sounds the more errors will result. The latter version of CAH holds that although errors cannot be predicted solely based on differences between the NL and TL, the errors that do occur in learner output can be explained based on these differences (Gass and Selinker, 2008). Because learner errors have been so central to understanding SLA as a process, it became necessary to distinguish between learner errors and mistakes. Learner errors as defined by Corder came to be thought of as systematic deviations in the learner's output of the target TL whereas mistakes were usually one-time occurrences in the learner's output that deviate from the TL. Once defined as such learner errors within a Cognitivist framework were equated with a learner's correct rule formation in his/her interlanguage (IL) where interlanguage is conceived as the learner's creation of a phonological system for the TL.

Eventually, as the field of SLA evolved, new concepts from phonology were introduced. One important concept that helped to establish more universally accepted cross-linguistic comparisons was markedness. Markedness is a concept that refers to a cross-linguistic description of how common or rare sounds are to many languages where those that are most common are unmarked and least common marked. Whereas in previous behaviorist frameworks learner difficulty was equated with differences between the NL and TL, in more recent frameworks that included markedness, it was hypothesized the more marked the feature the more difficult it would be. The Perceptual Assimilation Model (PAM), also more recently developed, has debated this view (Best, 1995). Within

PAM it is assumed that non-native speech sounds are *perceived* by learners in terms of their similarities and differences with native speech sounds. Predictions are made about how easily learners may discriminate between non-native contrasts based on how learners assimilate non-native sounds to native ones.

Consistent differences between the learner's output and the TL forms, traditionally referred to as errors, are described as phonological processes. While the traditional interpretation views the learner as lacking the knowledge of or capacity to produce a certain target-like feature, it is also possible to interpret the learner's role as an active one in which case phonological processes can be considered a way of simplifying the target sound as Ingram is cited in Hua and Dodd (2002). The processes include assimilation, deletion, and substitution. Assimilation refers to a transfer of sound features to an adjacent sound. Deletion refers to a sound not being produced. Substitution of sounds includes processes such as palatalization, rhotacization, velarization. Each substitution process refers to the articulator used in place of the target sound. When these processes involve movement of the sound toward the back of the oral cavity, they are referred to as backing. Other substitution processes are referred to by the manner of articulation involved in replacing the target sound. These include affrication, deaffrication, devoicing, deaspiration. Affrication typically involves inclusion of a stop in the segment whereas deaffrication involves deletion of the stop. Devoicing and deaspiration refer to the deletion of voicing and aspiration in the sound. Vowel replacement is a substitution process in which the target vowel is changed for another.

Evidence of the Difficulty of the Target Sounds

A number of researchers have discussed the difficulties learners seem to have with the dental, also referred to as the sibilant, palatal and retroflex groups of initial sounds. For the most part, these discussions have been based on observations of CFL learners or native speakers rather than on systematic studies of perception or production of these sounds. In line with CAH, Xing (2006) agreed with Norman's (1988) claim identifying the absence of the palatals and retroflexes in the learners' L1 as the reason for their difficulty. Norman justified this claim by observing that "(t)he English sounds 'j' [dʒ], 'ch' [tʃ], and 'sh' [ʃ] fall somewhere between the Chinese retroflexes and the palatals" (p. 140). Unlike Norman, Xing did not specifically address problems with /ʐʅ/ in the retroflex group, but included the sibilants as a difficult series, although only perceptually difficult for learners in the beginning. Supported by extensive observation through classroom teaching, Xing has concluded that the sibilants are relatively easy for learners yet require more explanation. One gap in her explanation is that it avoids addressing the role of the apical vowel and how this element impacts learners' perceptions and articulation of the dental consonants. This study proposes to help fill this gap via detection of learners' perceptual and production difficulties elicited on tasks targeting these sounds.

Relatedly, Dow (1972) highlighted certain dental sibilants, retroflexes, and palatals as sounds that cause confusion for English-speakers because they are uncommon or absent in their L1. Specifically, /tsɿ/ and /ts'ɿ/ are said to be confused in terms of aspiration. The English /ʃ/ is transferred for /ɕ/ and /ʂ/ and the English /dr-/, /tr-/, and /r-/ are used in place of the Chinese /tʂ-/, /tʂ'-/, and /ʐ-/ respectively. This finding seems to correspond

with the Perceptual Assimilation Model (PAM) that Best and others developed that proposes a crucial role for the L1 in the way non-native sounds are perceived (Sebastian-Galles, 2005). This would predict that adult learners, who are initially taught the difference between the L2 contrast, not able to fully perceive how the contrast is made, might deploy L1-based strategies in order to articulate a contrast, such as /tr-/ for /tʂ'/ and /palatal tʃ/ for /palatal tɕ'/. It would be helpful to determine whether these patterns will be borne out in the present study.

More recently, Zhu and Ye (1997) through classroom-based observations documented both perceptual and articulatory difficulties with the dental and retroflex initial consonants and the apical vowels that Chinese Second Language (CSL) learners of various L1 backgrounds studying abroad in China had, but also recommended content and methods for addressing those weaknesses. Shi and Wen's (2009) study focused on the production of Chinese vowels by American students and reported idiosyncratic and unstable articulation patterns for the apical Chinese vowels associated with dental and retroflex initial consonants, lending further support to PAM.

In addition to evidence of the difficulty of this set of sounds for second and foreign language learners, there is also evidence of their difficulty from first language acquisition. Hua and Dodd's (2000) reporting on the order of phonological acquisition showed that initial consonants generally stabilized last. Among the initial consonants, the palatals, a relatively marked or cross-linguistically uncommon sound set, emerged first and stabilized by two years of age. Although dentals and retroflexes emerged not long after other initial consonants, they tended to stabilize last. Specifically, the affricate dentals, 'zi' /tsɿ/ and 'ci' /ts'ɿ/, and fricative retroflexes, 'shi' /ʂʅ/ and 'ri' /ʐʅ/, were the last to stabilize between 4

and 4.5 years of age. The delay in stabilization of dentals and retroflexes was associated with the phonological process of fronting and backing. Hua and Dodd suggested cross-linguistic factors could have been at work in the language to which the children were exposed mitigating the effects of markedness or feature hierarchies that predict the palatal, dental, and retroflex groups would all emerge and stabilize last. Since blurring of the dentals and retroflexes is a fairly common feature identified with more southern varieties of Putonghua, such as the variety spoken in Shanghai (Chen, 1999), it is important to consider whether this apparent merging of syllables might occur in L2 learners as well.

Learning Difficulties Imposed by Chinese Writing Systems

Significantly, research has recently been conducted showing the direct role that orthography can play in the acquisition of L2 phonology. For example, Bassetti's (2007) research revealed how pronunciation problems still persisted even through the third year of Chinese study for Italian university students. Focusing on only three Chinese triphthongs /iou/, /uei/, and /uən/, Bassetti was able to link her participants' pronunciation errors to the deletion of the main vowel in the three triphthongs' Pinyin representations, '-iu', 'ui', and 'un'. In other words, what she demonstrated was a mismatch between the learners' mental phonological representations and the Pinyin representations causing a predictable pattern of learner errors in oral production, /iu/, /ui/, and /un/. Having established this link, Bassetti has helped to focus thinking on how instructors can address pronunciation problems that might otherwise go untreated. Her thoughtful analysis has also made it clear that popular claims of the transparency and simplicity of the Pinyin system may need more careful examination from teachers. Although it is indeed more transparent than English orthography, it is rather far from being a phonetic system of

transcription. While it might be obvious to some students that certain Pinyin letters represent more than one sound, one might suspect that when a single Pinyin letter, such as '-i' represents multiple L2 sounds absent in the L1, this could further complicate the acquisition of the sound by obscuring the relationship of sound to symbol. If the learning of new L2 sounds is a redeployment of knowledge of L1 features (Archibald, 2005), learners then might also be expected to go through a process of modifying certain L1 sound-symbol associations prior to being able to map new L2 sound-symbol associations derived from knowledge of L1 features. In addition to knowing the specific learning procedures required, it would also be helpful to know how long the learning process takes the typical learner so that teachers may develop more appropriate learning expectations and instructional techniques to address these issues more effectively.

In a comparative study of two groups of first-year university CFL students, Packard (1990) provides persuasive evidence of how a three-week time lag in the introduction of Chinese characters could result in meaningful differences in aural discrimination and transcription of unfamiliar Chinese monosyllabic words as well as in oral fluency. In addition, besides the lag group showing superiority to the no-lag group in phonologically related skills, significantly, the no-lag group did not consistently outperform the lag group in character writing or any other literacy skill at the end of the first or second semester of study. Such findings deserve careful consideration from teachers in terms of the impact instructional choices with regard to expectations for character learning can have on the development of phonologically related skills.

Relevant studies from Chinese linguistics, language pedagogy, and SLA, suggest that the attention the targeted groups of sounds have prompted from researchers, is due not

only to their difficulty for adult L2 and children L1 learners but also to the importance native Chinese speakers appear to place on them. Following Norman's observation that "the ability to distinguish the two series (the dentals and retroflexes) correctly has become one of the chief hallmarks of elegant standard (Chinese) pronunciation" (1988, p. 140), CFL instructors have sufficient reason to be concerned with their students' progress in acquiring these sounds. The social status that being able to discriminate these sounds can give speakers also suggests the potential stigma that may be associated with future speakers who might struggle with these sounds. In addition to sociocultural concerns, studies on the role of Pinyin and characters in CFL learning strongly suggests that developing perceptual acuity for these sounds provides a more solid linguistic foundation that can facilitate the process of acquiring more advanced literacy skills. Finally, a recent survey of the CFL field revealed an absence of studies on CFL pronunciation at the secondary level, as well as a need for research that can help identify effective instructional methods for pronunciation (Ke and Li, 2011), although "(i)t is essential for teachers to be fully aware of students' background and difficulty, and more importantly, find suitable ways to tackle students' difficulty and help them acquire pronunciation competence." (Xing, 2006, p. 90) To address these issues, the following research questions were developed:

Research Questions:

1. What patterns of error were detected among the target sounds that appeared to be resistant to acquisition?
2. Were the target sounds most difficult for learners to produce also those that were perceived by learners as being most difficult?

METHOD

Participants

The phonological acquisition of the proposed set of target sounds by 30 secondary-level CFL students was assessed. Students were recruited from two American public high schools, enrolled in three separate CFL classes, all of which were taught by the researcher and spanned the duration of the academic year from August 2010 through May 2011. Although participation in the study was entirely voluntary, all students were enrolled in the CFL course for academic credit. Six students, who were enrolled and attended the class regularly, opted out of the study. Based on self-reported information collected from each of the students regarding L1 or dominant language background and previous length of exposure to Mandarin, the students were divided into four groups for the analysis: Beginner Monolingual (BM), Beginner Asian Bilingual (BA), Non-beginner Monolingual (NM), and Heritage (H). 'Beginner' indicates no significant prior exposure to Mandarin was reported. 'Monolingual' indicates the dominant L1 background was reported as English. 'Asian bilingual' means the dominant L1 background reported included English and an Asian language other than Mandarin, specifically, Vietnamese, Laotian, and Hmong. 'Heritage' includes all learners who reported Mandarin was spoken in the home regardless of whether it was reported as the students' L1 or dominant language. (Polinsky and Kagan, 2007) It should be noted that five of the Heritage learners also reported Cantonese as spoken in the home and one among the five reported a third Asian language spoken in the home. This suggests there was significantly more linguistic diversity in the Heritage group than the label, 'Heritage' may capture. Participants' characteristics in each group are presented in Table 2 below.

Table 2. Characteristics of participant groups

Group	Ave. age	Male	Female	Total
Beginner Monolingual (BM)	15.3 yrs.	6	0	6
Non-beginner Monolingual (NM)	16.8 yrs.	4	1	5
Beginner Asian Bilingual (BA)	16 yrs.	3	2	5
Heritage (H)	16.3 yrs.	5	9	14
Total		18	12	30

Instruments

There were two instruments for this study, one measuring perceptions and the other production. The perceptual instrument included a survey of students' attitudes toward learning pronunciation and an assessment of students' perceived level of difficulty of and articulation of the target sounds. The survey of students' attitudes and beliefs about pronunciation learning consisted of seven statements reflecting opinions related to pronunciation learning to which students responded using a Likert-type scale ranging from one to five depending on how strongly they agreed or disagreed with each statement. See Appendix A for Survey of Attitudes toward Learning Pronunciation. The assessment of students' perceived level of difficulty of the target sounds, intended to elicit how students implicitly conceptualized the target sounds, consisted of four parts. Each part asked students to perform specific tasks. See Appendix B for Assessment of Learners' Perceptions toward the Target Sounds. The first task, designed to elicit knowledge of the three vowel-types as rhyming sounds, was to categorize the target syllables based on the likeness of their vowel sounds and then to rank or identify how difficult they felt these

sounds were in relation to each other. The second task, designed to elicit strategies for producing differences between target sounds, was to explain how to produce the difference between target syllables from the dental and retroflex groups. The third task, designed to elicit L1 sound transfer or redeployment, prompted students to suggest English words that seemed to them to contain or approximate the target syllables from the retroflex group. The fourth task, designed to elicit detection of articulatory locus of consonant sounds, prompted learners to indicate the place of articulation on a diagram of the mouth for three syllables, each from separate groups.

The pronunciation assessment was an oral reading task, composed of two sections, intended to elicit students' proceduralized knowledge of the target sounds. To elicit this knowledge, conditions were simulated to reflect how students might rehearse new vocabulary typically introduced at the beginning of a new lesson in their textbook. The first section listed 55 monosyllabic syllables providing six samples of each of the ten target syllable types and included at least one sample of each of the four tones, wherever possible. Items were grouped by syllable-type so that no more that six target syllables of a particular syllable-type were sequenced together. In constructing the list, *the Third edition Pocket Oxford Chinese Dictionary (Manser, 2004)* was referenced to ensure accuracy and currency of expressions, as much as reasonably possible, and a Chinese native speaker helped with proofreading. Eighty percent of the characters and phrases were those to which students had been exposed in a first year course. The second section listed 60 disyllabic phrases or words that were generated from the samples on the first list and formatted similarly to the monosyllabic list. An effort was made to ensure that all target syllable types occurred in first and second position of the disyllabic phrases an equal number of times, wherever

feasible. See Appendix C for Pronunciation Assessment. Interrater reliability for both sections of the pronunciation assessment were calculated based on five participants' scores as the Pearson product moment correlation coefficient using Excel software. The Pearson coefficients were 0.968 and 0.972 for the monosyllabic and disyllabic sections, respectively, suggesting a high degree of agreement between raters for both sections of the assessment.

Procedures

Approval for research was received from the Human Subjects Committee Lawrence Campus (HSCL) under an expedited procedure waiving the requirement for signed parental permission. No further permission from the cooperating institution was necessary considering data was collected during regular class time. In order to minimize the impact of the attitude survey on the students' performance on the pronunciation assessment, the attitude survey was administered a week prior to the pronunciation assessment. Since the implicit purpose of the survey was to raise awareness of the importance of pronunciation to language learning, students were encouraged to talk to each other about the survey only after everyone had completed it. No input from the teacher/researcher was provided to the students during the survey other than attempts to satisfy their questions. The students' pronunciation was assessed and audio-recorded individually during the last two weeks of the school year. QuickTime Player 7.6.9 Audio Recording Software on a Mac OS X MacBook Pro with internal microphone was used for the recording. Students were instructed how to record themselves after testing the equipment, shown how to monitor the volume of their input on the device, asked to read each syllable or phrase only one time each, starting at the top of each page and reading to the bottom of the page one line at a time before moving to the next page, and directed to read clearly and at a comfortable pace without assistance.

Before beginning to record, students were informed they would be able to complete the recording in three to five minutes with the understanding that there was no time limit.

Scoring and Transcription

Each consonant, vowel, and tone of each target syllable in each item of the recording was scored in terms of accuracy on a three-point scale based on target-likeness, a score of '1' representing dissimilarity to the target sound, a '2' representing somewhat similar but distracting to communication, and a '3' representing very similar and easily understood by a native speaker. See Appendix D for Instructions for Rating Production Accuracy of Target Syllables with Final '–i'. The researcher transcribed each sound receiving a '1' using IPA in order to identify how the sound deviated from the target. These transcriptions were also referred to in identifying the phonological processes that appeared in the learners' interlanguage.

Data Analysis

Following Hua and Dodd (2000), differences in production of the same phoneme in the same syllable position by different L1 speakers can reflect acquisition difficulties. In their study on L1 acquisition, stabilization of a sound was defined at 66.7% accuracy for 90% of the children in each age group. Because the pronunciation assessment used in the present study relied on a three-point scale as a measure of accuracy, where 66.7% is the equivalent of 'somewhat target-like, but possibly distracting to a listener', a somewhat higher score of 75% for each sound was set as a reasonable target for classroom learners, considering 25 of the 30 participants agreed on the survey of learning attitudes that a native-like pronunciation was an important learning goal of theirs. See Appendix E for Results of CFL Learner Survey on Attitudes toward Pronunciation.

RESULTS AND FINDINGS

Difficult-to-produce Sounds

Among the target sounds assessed, the most difficult sound for learners on the oral production assessment as shown in Appendix F was 'ci' /ts'ɿ/. Overall it was the only sound that fell below the 75% accuracy rating for all groups for which the rating was 67.3%. The consonant and vowel scores of the two beginner groups most clearly reflected difficulties taking place with 'ci' /ts'ɿ/. Group BA accuracy ratings averaged 43.3% on the consonant 'c' /ts'/, while Group BM averaged 59.3% for the same sound. In addition, the same groups' vowel accuracy for 'i' /ɿ/ reached 57.8% and 68.1%, respectively. There was perhaps only one wrinkle in the Group H's performance on the monosyllabic section of the assessment with an almost borderline average accuracy rating of 74.2% on the vowel sound, 'i'/ɿ/, in 'ci'.

Sounds other than 'ci' /ts'ɿ/ that appeared to present learners with difficulty did so only for Groups BM and BA. Group BM appeared to reflect difficulty with the apical vowel, /ɿ/, in 'si' and the palatal consonant, /ɕ/, in 'xi' on the monosyllabic section, averaging 70.4% and 74.1% accuracy, respectively. In Group BA, the consonant sound, /ʐ/, in 'ri' seemed to present greater difficulty for learners, with average accuracy ratings reaching only 60.6%. Difficulty with the 'ri'-vowel, /ʅ/, seemed to accompany difficulty with the 'ri'-consonant, /ʐ/, the average accuracy of which was nearly borderline at 73.9%.

Five other sounds that seemed to present minor difficulty for Group BA included the dental vowel, /ɿ/, in 'si' and 'zi', as well as the consonants /ɕ/, /s/, and /tʂ/in 'xi', 'si', and 'zhi', respectively, and the high front vowel, /i/, in 'xi'. As in Group BM, accuracy with these

sounds was more likely to increase when produced as a part of a phrase. For example, Group BM average scores on monosyllabic items targeting 'ci' /tsʰɿ/ increased from 64.8% to 71.3% on disyllabic items, and on 'si' /sɿ/ increased from 70.4% to 77.8%. Similarly, Group BA average scores on monosyllabic items targeting 'ci' /tsʰɿ/ increased from 53.3% to 62.2%, and on 'si' /sɿ/ increased from 63.3% to 81.1%. One case reflecting this pattern was found in Student Y from Group BA, who consistently produced [ə] for the vowel in 'si' /sɿ/ on monosyllabic items, but produced a target-like [ɿ] on its first four disyllabic items.

Patterns of Errors in Difficult Sounds

Following Hua and Dodd (2000), patterns of errors in the production of difficult sounds were analyzed and labeled in terms of phonological processes. The two sounds that were targeted for analysis were 'ci' /tsʰɿ/and 'ri' /ʐɿ/ because learners across all four groups reflected difficulties production-wise with both of these sounds. The phonological processes affected initial consonant sounds and final vowels in these two syllables relating to assimilation, deletion and substitution. The most frequently occurring processes for each of these sounds are detailed below.

For 'ci' the most commonly occurring phonological process was deaffrication. Deaffrication took place for this sound in disyllabic phrases in either the first or the second syllable in at least 12 learners. For example, 'ci' /tsʻɹ/ was realized as 'si' [sɹ] in '生詞' 'sheng-1ci-2' (vocabulary) and '瓷器' 'ci-2 qi-4' (porcelain), among others. Backing of 'ci' /tsʻɹ/ occurred as palatalization in four learners, rhotacization in three learners, and velarization in one learner. Backing is where the articulatory locus is moved further back and is reported by Hua and Dodd (2000) as one of the most frequently occurring processes

in Chinese-speaking children. Examples include 'ci' /ts'ɹ/ rhotacized as [tʂ'ɻ] in '瓷器' 'ci-2 qi-4' (porcelain), palatalized as [tɕ'i] in '磁鐵' 'ci-2 tie-3' (magnet), and velarized as [kɤ] in '三次' 'san-1 ci-4' (three times). These processes occurring in 'ci' were found only in Groups H and BA. Deaspiration of 'ci' /ts'ɹ/ took place in at least three different learners. Deaspiration only appeared for 'ci' /ts'ɹ/ in learners from the Groups BM and BA. As an example, 'ci' /ts'ɹ/ was realized as [tsai] in '刺耳' 'ci-4 er-3' (screechy). Vowel replacement in the 'ci'-syllable was one of the more common processes that took place among all learner groups in various ways, including [i], [ai], [-ɻ], [ɛ], [ɤ], [ə] for /-ɹ/.

The phonological processes resulting in the error patterns for 'ri' were sound deletion, affrication, devoicing, palatalization, and vowel replacement. Sound deletion was the most prevalent and it was the only process that the heritage group displayed for 'ri' /ʐɻ/, producing [əɹ] instead, losing its initial consonant sound without causing communicative breakdowns. Affrication and devoicing for 'ri' /ʐɻ/resulted in [tʂɻ] for at least three different learners. Vowel replacement was by far a less active process for 'ri' /ʐɻ/than it was for 'ci'/tsʾɻ/. For 'ri' /ʐɻ/, there were only two substitutions: the central vowel [ɤ] and the high vowel [i] for /ɻ/.

Sounds Perceived as Difficult

There was some consensus around which sounds were more difficult. Patterns of response indicating perhaps why that was so were also found. A majority of learners ranked dentals as the most difficult sounds in comparison to palatals and retroflexes. 11 of 16 of those learners specifically named 'ci' /tsʾɻ/ as the most difficult syllable among the ten assessed. In addition, although the fewest number considered retroflexes most

difficult, ten learners identified ‘ri’ /ʐɻ/ as most difficult. Results are presented in Appendix G.

The degree of difficulty learners perceived for ‘ci’ /tsʾɿ/ were manifested on both the syllable-grouping task and diagramming task. On the syllable-grouping task, 10 of 15 learners misidentified dental syllables with palatals. 9 of these 10 learners misidentified ‘ci’ /tsʾɿ/ with other palatal syllables, suggesting that learners confused the apical vowel ‘i’ /ɿ/ with the high front vowel ‘i’ /i/, both represented by the same Pinyin spelling, ‘-i’. Results are shown in Appendix H. On the diagramming task, students had relatively less success in mapping the locations of ‘ci’ /tsʾɿ/, ‘qi’ / tɕʾi/, and ‘chi’ /tʂʾɻ/ than they did grouping the syllables according to their rhyming sounds. That is, while 15 learners successfully categorized syllables into three groups based on the likeness of their vowel sounds, only ten students located ‘ci’ /tsʾɿ/, ‘qi’ / tɕʾi/, and ‘chi’ /tʂʾɻ/ on the diagram in a way that showed a front-to-back relation among the consonant sounds. A similar number of learners located these sounds in nearly the same place on either the blade of the tongue and/or the alveolar ridge among other loci. Results are presented below in Appendix I.

With regard to ‘ri’ /ʐɻ/, perceived difficulties were observed on the syllable-grouping task and the L1 sound-transfer task. Although only ten learners specifically named ‘ri’ /ʐɻ/ as a difficult sound, of the 12 learners who misidentified retroflex syllables with dentals or palatals, 11 of them misidentified ‘ri’ /ʐɻ/, eight having done so with dentals and the other three with palatals. On the L1 sound-transfer task, responses varied among five types. The most common type was a non-response where 10 of 30 learners left the item blank or suggested they were not sure. For example, one learner claimed he

"could not think of anything for the ideal 'r' sound in Chinese." The other two-thirds of the group, however, did suggest an English equivalent for 'ri' /ʐʅ/. Suggested equivalents included both voiced and voiceless English affricates, such as the 'jer' in 'Jergens' and the 'ture' in 'puncture'. Voiced and voiceless fricative equivalents included the '-sure' in 'measure', the 'shir-' in 'shirt', and the 'sir' in 'sir'. Other equivalents suggested were the final 'r' and the digraph 'dr'. Of the 20 responses only 11 of the suggested equivalents included English retroflex vowels. Results are provided in Appendix J.

Mismatches in Production and Perceived Difficulty

Mismatches were found in the results for production and perceptual difficulty. Eight learners identified the palatals as the most difficult sounds. However, only half of these learners showed difficulties with palatals in the production assessment. The other half obtained a high degree of accuracy on the production assessment, although they still perceived the palatals as most difficult. Representative of the latter type of learner, Student X from Group BM was one who not only correctly grouped the ten syllables according to their rhyming sounds, but also identified the palatal group as the most difficult, writing that 'xi' and 'qi' were "troublesome to find the correct sound". Based on Student X's production scores alone, it would appear as though Student X had little if any difficulty with the palatal group, receiving mostly scores of '3' for all three syllables on monosyllabic and disyllabic sections, suggesting that Student X and his peers' concern for the palatals was unwarranted. Nevertheless, it is also worth noting that on the diagramming task, Student X incorrectly identified the 'qi' /tɕʰi/ sound as being produced by "both sets of teeth" and indicated that 'ci' /tsʰʅ/ and 'chi' /tʂʰʅ/were both located at the lower teeth as well.

DISCUSSION AND IMPLICATIONS

Production Difficulties

The results obtained reflecting learners' production difficulties with the consonant sound in 'ci' /tsʾɿ/ differ with Xing's observation that the dentals are relatively easier for English speakers. This finding calls into question the adequacy of a weak version of CAH used to explain L2 sounds already existing in a learner's L1 system as reason for their not presenting difficulties for the learner. On the other hand, results demonstrating learner difficulties with the apical vowel sound, 'i' /ɹ/, in 'ci' /tsʾɿ/ seem to lend limited support for CAH, if the notion that 'ci' /tsʾɿ/ contains an apical vowel is accepted.

Interestingly, 'ri' /ʐɻ/was the only retroflex vowel with which Groups BA reflected persistent patterns of difficulty, suggesting the demands of the 'ri'-consonant, /ʐ̩/, on attentional resources might have contributed to momentary destabilizing of the vowel in some learners. Although it could be argued that learners might produce more accurate representations of this sound on tasks emphasizing the conveyance of meaning over form, the increase in task demands on the learner could still result in similar issues of accuracy that were seen in many learners on this assessment.

The results showing improved accuracy of target sounds in disyllabic phrases over those in monosyllabic words suggested that contextual cues could have been playing a role in activating learners' phonological representations of the apical vowel, 'i' /ɿ/ (Bassetti, 2007), otherwise obscured in isolation or unfamiliar phrases by the spelling's correspondence with several different values, e.g. /i/, /- ɹ /, /- ɻ /, and /-ei/ as in 'dui'. If this interpretation is justified, then teachers should expect learners' knowledge of sound-

symbol relationships to develop over time, facilitated by ample amounts of contextualized input that is comprehensible to the learner (Krashen, 1981), as occurred with the two non-beginner groups who showed greater accuracy in producing 'ci' /tsʾɿ/, suggesting growth had occurred over time with greater exposure and opportunity for practice and feedback.

Phonological Processes

Deaffrication also is known to occur in English for the Japanese loan-word 'tsunami', which is often pronounced as [su na' mi] and widely accepted as such presumably because English phonology does not permit aspiration of alveolar affricates in initial position (Merriam-Webster, 2012). If deaffrication is related to L1 transfer, it may help to explain the deaffrication of 'ci' /tsʾɿ / in 'ci-2 qi-4', but not 'sheng-1 ci-2', where the alveolar affricate occurs in mid-position, something that is permitted in English, mostly for foreign loan words such as 'pizza', 'barmitzvah', and 'Mitsubishi' or at the phrase-level with elision practices such as 'Cats are finicky.' and 'She hates it'. Notwithstanding, alveolar affricates following velar nasals as in 'sheng-1 ci-2' are perhaps relatively more rare in English, e.g. 'He yanked some rope.' If these factors suggest L1 transfer is not sufficiently effective for phonological acquisition, it may make more sense for teachers to explore other strategies available to learners that target how students conceptualize L2 sound patterns independently of the L1. See Appendix F for Strategies Used by Learners in Distinguishing 'zi' and 'zhi'.

The variety of representations produced for the final vowels in 'ci' /tsʾɿ/ and 'ri' /ʐɻ/shown in the results suggest rather strongly that many if not most of these substitutions are more likely a product of L2 sound system-building that learners are

attempting in constructing their interlanguage primarily through hypothesis testing of sounds based on the accumulation of L2 intake over time rather than as a product of L1 sound transfer alone. It is remarkable, for example, that while vowel replacement was far less active for 'ri' /ʐɿ/ than for 'ci' /tsʰɿ/, 'ri' /ʐɿ/ and 'ci' /tsʰɿ/ at least shared the same substitutions, suggesting that this process at times was also motivated by misidentification of 'ri' /ʐɿ/ with the dental group. The relative obscurity of the '-i' for many learners due to its correspondence with many sounds in Chinese should not be underestimated though, as having the potential for motivating learners to engage in deeper analysis of sound-symbol relationships. If teachers ignore such challenges, important opportunities for teachable moments may be wasted.

Perceived Difficulty

Learner's perceived difficulty of 'ci' /tsʰɿ/ and 'ri' /ʐɿ/ corresponded with certain difficulties exposed on the production assessment. For example, misidentification of vowel sounds corresponded with the replacement of apical vowels /ɿ/for high vowels /i/. Backing of 'ci' /tsʰɿ/ and affrication of 'ri' /ʐɿ/ on the production assessment also corresponded with perceived difficulties related to identification of their consonant sounds on perceptual tasks. That these correspondences between perceived difficulty and production difficulty exist suggests not only that perceptions are important factors in production, but seem to also reflect areas in the learners' developing interlanguage system that are still subject to learners' hypothesis testing. Considering the interaction that may occur between perception and production, it is suggested that teachers ought to look at these processes as important learning assets, instead of merely as errors or deficiencies,

upon which conceptual and articulatory strategies can allow learners to continue developing ways of refining and distinguishing more effectively between target sounds that are easily conflated.

In other words, CFL teachers in secondary classrooms should expect Chinese segmental sound acquisition to be an on-going process for all types of learners; therefore, constant sound hypothesis testing should be encouraged. Since learners are likely to have and recognize pronunciation weaknesses with 'ci' /tsʲɻ/ and 'ri' /ʐɻ/ at the end of a year or more, without necessarily knowing why, it is suggested that teachers highlight these sounds through extensive aural discrimination and identification tasks emphasizing perceptual familiarity combined with oral and written textual enhancement strategies as soon as learning starts in order to increase exposure to, raise learners' awareness of, and engage their self-monitoring ability for these sounds (Han et al., 2008). This may include emboldening, underlining, circling, italicizing, or color-coding printed input or slowing down, repeating, or emphasizing with salient intonation in oral or non-verbal input. If self-regulated learning is the ultimate goal of teaching, then it is as critical to train learners' in detecting their own pronunciation difficulties as it is to provide correct models and explanations of how to produce these sounds correctly. To this end, spectrography, a technology providing visual feedback of learner output, allows learners to monitor their output by enabling them to compare their own output with native-like output (M. Gonzalez-Bueno, personal communication, February 22, 2013).

While the strategy of L1 sound transfer can serve as a useful starting point for, and is sometimes not only inevitable, but also preferred by many learners, this is not to suggest that it is the best or most useful strategy for all learners to start with. Considering that one

of the main perceptual paths through which literate learners access sounds on their own is through an alphabetic Pinyin system, learners' natural tendency to rely on L1 sound-symbol transfer as a sort of default mode cannot be underestimated. Often times, CFL teachers assume that the shared alphabetic system for English and Chinese is an advantage to American learners. However, if learners are simultaneously required to decode two writing systems, i.e. Pinyin and characters, teachers should also ask how this unique burden on the learner will impact the sound acquisition process and more generally the L2 acquisition process.

Implications of Divergences in Perceived Difficulty and Production

The evidence from Student X's case suggests that conceptually Student X had not detected a front-to-back relationship among the three consonant sounds in 'ci'/tsʾʅ/, 'qi' /tɕʾi/, and 'chi' /tʂʾʅ/, something Hua and Dodd (2000) considered a crucial distinction in the acquisition of these sounds. If this is interpreted to mean that Student X lacked a clear sense of where 'qi' /tɕʾi/ is located in relation to 'ci' /tsʾʅ/ and 'chi' /tʂʾʅ/, the proximity of 'qi' /tɕʾi/ to the other sounds makes it conceptually difficult for learners, although difficulty conceptualizing the sound still may not entail difficulty producing it. That is, the palatals were ranked as the most difficult sounds, but while no difficulty was reflected production-wise for the palatals, locating 'qi' /tɕʾi/ in relation to 'ci' /tsʾʅ/ and 'chi' /tʂʾʅ/ on the diagram of the mouth was performed less effectively. Given the divergences between learner perceived difficulty and production, it should be recognized that production scores alone may not generate an adequate picture of which L2 sounds learners really know. In other words, if teachers rely too heavily on production scores for deciding

on learner feedback, they may unwittingly neglect important aspects of sound acquisition more directly relevant to individual learning needs, such as aural discrimination skills and the ability to relate symbol to sound and vice-versa.

Suggested Instructional Strategies

In light of the results of this study, textually enhanced input has already been suggested as one promising strategy used to raise noticing and awareness among learners of the targeted sounds. That combined with appropriately timed individualized corrective feedback are thought to be effective ways of promoting acquisition of particular forms in a task-based/ meaning-based classroom environment. Current trends in task-based approaches to foreign language teaching strongly suggest meaning-oriented tasks versus more traditionally decontextualized exercises. Specific examples include 'dictogloss' or essentially what are dictation-type tasks where learners are asked to listen to a message conveyed either orally or in print and then attempt to reconstruct the message either individually or in pairs or small groups. It is conceivable that such messages could be specifically designed by teachers or textbook designers to focus on the targeted forms. Such tasks, providing learners with opportunity to hear natural oral input of the language, also have the potential of engaging learners in producing the target forms in both speech and writing, thus creating the conditions necessary for learners to relate sound to symbol in a meaningful context. Opportunities for learners to check, discuss, and revise their work may increase the likelihood over time of acquisition of these sounds. This in concert with teachers' feedback and seizing on teachable moments to focus on form are possibly effective tools worth further investigation.

Limitations and Future Research Directions

There were a number of limitations to this study that if addressed could add to the accuracy and significance of future research. First of all, the bulk of the data collected and analyzed in this study, as noted above, was from a reading-aloud task rather than from spontaneous speech. While the controlled conditions of the reading-aloud task facilitated collection of the sounds targeted in the study, data collection under more natural conditions would possibly strengthen its validity as an acquisition study. Secondly, as a cross-sectional study, the sample size of the groups was a concern. Alternatively, a future study could perhaps benefit more from a longitudinal approach. Finally, this study primarily measured learners' production of difficult sounds with only limited evidence of learners' perceptions toward the sounds, specifically their perceptions of the degree of difficulty and articulation of the sounds. It is believed that measuring learners' aural perceptions of these sounds could shed more light on whether and at what stages learners experience difficulty discriminating or identifying difficult sounds. It is suggested that future research should aim in this direction to help determine the appropriate timing of listening versus production tasks for classroom pronunciation instruction.

CONCLUSION

This study demonstrated that group and individual data on secondary CFL learners' attitudes, perceptions and productions of target sounds provide important insights into how learners come to discriminate among closely related sounds with a focus on the most difficult sounds, 'ci' and 'ri'. Measurements of four groups' perceived and productive difficulties presented evidence supporting that L2 learners' phonological development reflects patterns similar to Hua and Dodd's L1 phonological acquisition study of Chinese

children. A striking similarity was that the aspirated dental affricate, 'ci', continued to present challenges to L2 learners at the end of a year long Chinese course. This finding lends support to PAM while challenging the CAH that explains difficulty in terms of the differences between L1 and L2 sounds alone.

This study also elicited processes and strategies that individual learners engage in on a variety of discriminatory tasks. Learners employed strategies that included but not limited to ranking sounds by difficulty, L1 sound transfer, distinguishing syllables by rhyme, relating sound to articulatory locus, identifying articulatory organ, and distinguishing sounds by quality or manner. By assessing the frequency and success with which learners employ these strategies, teachers simultaneously raise learner awareness of problematic L2 sounds as well as expose learning resources that can be exploited more strategically to effectively target individual learning needs. While it was suggested, however, that some L2 learners will eventually acquire difficult sounds along with how they are encoded graphically, the fact that learners may still perceive difficulties with these sounds even after they are able to produce them on assessments suggests that teachers should give further serious consideration to instructional practices that promote meaningful acquisition of L2 sounds in Chinese along with L2 sound system building that is increasingly independent of L1 transfer.

REFERENCES

Archibald, J. (2005). Second language phonology as redeployment of L1 phonological knowledge. The Canadian Journal Linguistics, 50, 285-314.

Bassetti, B. (2007). Effects of hanyu pinyin on pronunciation in learners of Chinese as a foreign language. London: Birkbeck ePrints.

Best, C. T. (1995). A direct realist view of cross-language speech perception: New directions in research and theory. In Winifred Strange. *Speech perception and linguistic experience: Theoretical and methodological issues*. Baltimore: York Press. pp. 171–204.

Chen, P. (1999) *Modern Chinese: History & Sociolinguistics.* New York: Cambridge University Press. (e-book)

Dow, F. D. M. (1972). *Outline of Mandarin Phonetics* (2nd ed.). Canberra: Australian National University Press.

Duanmu, S. (2007). *The Phonology of Standard Chinese (2nd ed.).* Oxford: Oxford University Press.

Gass, S. and Selinker, L. (2008). *Second Language Acquisition: An Introductory Course* (3rd ed.). New Jersey: Routedge. (e-book)

Hartman, L. M. (1944) The Segmental Phonemes of the Peiping Dialect. *Language*. 20:1, 28-42.

Han, Z-H et al. (2008). Textual enhancement of input: Issues and possibilities. *Applied Linguistics*. 29: 597-618.

Howie, J. M. (1976). *Acoustical studies of Mandarin vowels and tones.* Cambridge: Cambridge University Press.

Hua, Z. and Dodd, B. (2000). The phonological acquisition of Putonghua (Modern Standard Chinese)*. *Journal of Child Language.* 27, 3-42.

Ke, C. and Li, A. (2011). Chinese as a foreign language in the US. *Journal of Chinese Linguistics*. 39:1, 177-238.

Krashen, S. (1981). *Second language acquisition and second language learning.* Oxford: Pergamon Press.

Manser, M. H. et al. (eds.) Pocket Oxford Chinese Dictionary: English-

Chinese, Chinese-English, 3rd ed. (2004). USA: Oxford University Press.

Norman, Jerry. (1988) *Chinese.* Cambridge: Cambridge University Press.

Packard, J. L. (1990). Effects of time lag in the introduction of characters into the Chinese language curriculum. The Modern Language Journal, 74: 2, 167-175.

Polinsky, M. and Kagan, O. (2007). Heritage languages: In the 'wild' and in the classroom. *Language and Linguistic Compass*. 1:5, 368-395.

Sebastian-Galles, N. (2005). Cross-language speech perception. In D. Pisoni and R. Remez (eds.) The Handbook of Speech Perception. (pp. 546 – 566) Oxford: Blackwell Publishing.

Shi, F. and Wen, B. (2009) 中美学生元音发音中的母语迁移现象研究（Zhongmei xuesheng yuanyin fayinzhongde muyu qianyi xianxiang) [Study of Language Transfer in Vowel Articulation by Chinese and American Students]. *Journal of the Chinese Language Teachers Association,* 42:2, 17-32.

Sun, C. (2006). *Chinese: A Linguistic Introduction.* Cambridge: Cambridge University Press.

Thompson, S. and Li, C. (1987). Chinese. In B. Comrie (ed.) *The World's Major Languages*. (pp. 811-833) New York: Oxford University Press.

Tsunami. (2012). *Merriam-Webster.com.*Merriam-Webster, Inc. Retrieved August 1, 2012, from http://www.merriam-webster.com/dictionary/tsunami.

Xing, Janet Zhiqun. (2006) Teaching and Learning Chinese as a Foreign Language: A Pedagogical Grammar. Hong Kong. Hong Kong University Press.

Zhu, C., Ye, Jun, Cui, L, and Lin, X. (Eds.) (1997) *Waiguo xuesheng hanyu yuyin xuexi duice* [Studies of and Strategies for Chinese Foreign Language Learner Pronunciation, trans. Ring (or Chinese phonetics for foreign students, trans. Bassetti)]. Beijing, China: Yuwen.

APPENDIX A. Survey of Attitudes toward Pronunciation

Based on your own opinion, how strongly do you agree with the following statements:
1= I strongly disagree. *2= I somewhat disagree.* *3= I have no opinion.*
4= I somewhat agree. *5= I strongly agree.*

1. Pronunciation enhances self-image.

   1 2 3 4 5

2. Pronunciation enhances social relationships.

   1 2 3 4 5

3. Pronunciation can be acquired easily.

   1 2 3 4 5

4. One's pronunciation is determined by birth, origin, or mother tongue.

   1 2 3 4 5

5. Learning pronunciation is fun.

   1 2 3 4 5

6. I consider learning a native-like or near native-like pronunciation an important goal of mine.

   1 2 3 4 5

7. I feel comfortable speaking the target language in front of my peers.

   1 2 3 4 5

APPENDIX B. Assessment of Learners' Perceptions toward the Target Sounds

Task 1. Based on your own understanding, please answer the following as fully as possible:

1. The following Chinese syllables all contain the same vowel, '-i'. Arrange the following syllables into three groups: Group A, Group B, and Group C, *based on the likeness of their vowel sounds*: zi, zhi, ji, ci, chi, qi, si, shi, xi, ri

Group A.	Group B.	Group C.

1.1 Of the three groups of sounds above, how would you rank them in order of difficulty?
Most Difficult:
Not as Difficult:
Least Difficult:

1.2 Among the Most Difficult Group, in your opinion, is one or more sound particularly difficult?

1.3 If so, which one or which ones?

Task 2. In the space below, please describe in your own words to a novice learner how to produce the difference between 'zi' and 'zhi'?

Task 3. What English words, if any, would you suggest contain the following sounds?

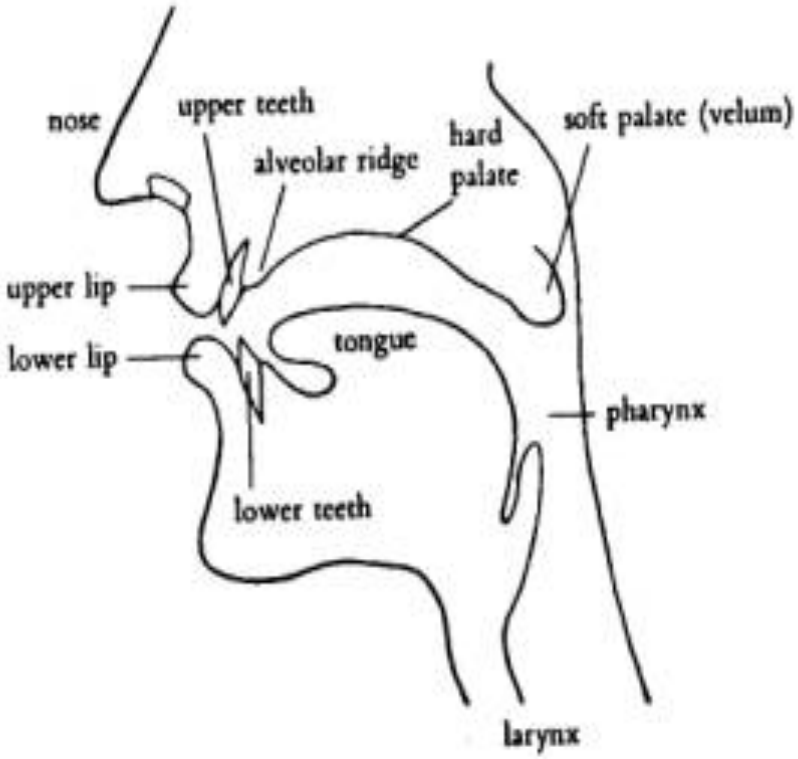
Chi: _____

Shi: _____

Ri: _____

Task 4. Draw arrows on the diagram pointing to the nearest locations in the mouth where you feel the sounds, 'ci', 'chi', and 'qi', are produced.

Diagram of Speech Anatomy



APPENDIX C. Pronunciation Assessment

Target Syllables in Isolation

A1	zǐ	子	child
A2	zì	字	character or word
A3	zī	资	capital, money
A4	zǐ	紫	purple
A5	zì	自	self
A6	zī	姿	posture
B1	zhī	支	to pay
B2	zhī	知	to know
B3	zhí	直	straight
B4	zhǐ	只	only
B5	zhǐ	纸	paper
B6	zhǐ	指	to point
C1	jǐ	几	how many?
C2	jī	机	machine
C3	jī	鸡	chicken
C4	jì	记	to remember
C5	jí	急	impatient
C6	jí	级	grade, level
A1.1	cí	词	phrase
A2.1	cì	次	time
A3.1	cí	瓷	porcelain
A4.1	cì	刺	pierce
A5.1	cí	磁	magnet
A6.1	cí	辞	to depart
B1.1	chī	吃	to eat
B2.1	chǐ	齿	teeth
B3.1	chí	池	pond
B4.1	chǐ	尺	foot (measurement)
B5.1	chí	匙	spoon
B6.1	chì	赤	red
C1.1	qī	七	seven
C2.1	qí	奇	strange
C3.1	qì	气	air
C4.1	qí	其	his, her, its
C5.1	qī	期	period of time
C6.1	qǐ	起	to get up
A1.2	sì	四	four

A2.2	sǐ	死	to die
A3.2	sī	思	to think
A4.2	sī	司	to have charge of
A5.2	sī	丝	silk
A6.2	sī	斯	this; then
B1.2	shí	十	10
B2.2	shī	师	master
B3.2	shǐ	始	to begin
B4.2	shì	事	an affair or matter
B5.2	shì	室	room
B6.2	shì	识	to recognize
C1.2	xī	西	west
C2.2	xǐ	洗	to wash
C3.2	xǐ	喜	joy
C4.2	xī	希	to hope
C5.2	xí	习	to study
C6.2	xì	戏	performance
B1.3	rì	日	day

Target syllables in Phrases

A1	hái zǐ	孩子	child
A2	hàn zì	汉字	Chinese character
A3	gōng zī	工资	wages
A4	zǐ sè	紫色	purple
A5	zì jǐ	自己	self
A6	zī shì	姿势	posture
B1	zhī piào	支票	a check
B2	yù zhī	预知	to know beforehand
B3	zhí zǒu	直走	go straight ahead
B4	bù zhǐ	不只	not only that
B5	bái zhǐ	白纸	blank paper
B6	zhǐ dìng	指定	to appoint
C1	jǐ gè/ge	几个	how many?
C2	fēi jī	飞机	airplane
C3	jī ròu	鸡肉	chicken
C4	wàng jì	忘记	to forget
C5	jí shì	急事	an urgent matter
C6	nián jí	年级	grade in school
A1.1	shēng cí	生词	vocabulary
A2.1	sān cì	三次	three times

A3.1	cí qì	瓷器	porcelain
A4.1	cì ěr	刺耳	screechy
A5.1	cí tiě	磁铁	a magnet
A6.1	gào cí	告辞	to say good-bye
B1.1	chī fàn	吃饭	to eat
B2.1	yá chǐ	牙齿	teeth
B3.1	chí táng	池塘	a pond
B4.1	yīng chǐ	英尺	an English foot
B5.1	tāng chí	汤匙	a spoon
B6.1	chì jiǎo	赤脚	bare-footed
C1.1	shí qī	十七	seventeen
C2.1	qí guài	奇怪	strange
C3.1	shēng qì	生气	to be angry
C4.1	qí shí	其实	actually
C5.1	xīng qī	星期	week
C6.1	qǐ chuáng	起床	to get up from bed
A1.2	sì shí	四十	forty
A2.2	shēng sǐ	生死	life and death
A3.2	yì sī	意思	meaning
A4.2	gōng sī	公司	a company
A5.2	sī chóu	丝绸	silk
A6.2	sī wén	斯文	cultured and refined
B1.2	shí wàn	十万	100,000
B2.2	lǎo shī	老师	teacher
B3.2	kāi shǐ	开始	to begin
B4.2	shì qíng	事情	an affair or matter
B5.2	jiào shì	教室	classroom
B6.2	shì zì	识字	to be literate
C1.2	dōng xī	东西	a thing
C2.2	xǐ shǒu	洗手	to wash hands
C3.2	xǐ huān	喜欢	to like
C4.2	xī wàng	希望	to hope
C5.2	fù xí	复习	to review
C6.2	yóu xì	游戏	a game
B1.3	rì lì	日历	calendar
B1.3.1	jié rì	节日	a festival
B1.3.2	rì wén	日文	Japanese language
B1.3.3	rì jì	日记	a journal or diary
B1.3.4	míng rì	明日	tomorrow
B1.3.5	shēng rì	生日	birthday

APPENDIX D. Instructions for Rating Production Accuracy of Target Syllables with Final –i

Preliminaries of Study: The recordings are of first-year Chinese Foreign Language Learners. The target language of the recordings is Standard Mandarin Chinese, known as Putonghua. All subjects are sight-reading from a word list without prior rehearsal. Their instructions were to give a natural reading of each item on the list in the order of the list, one at a time, one time each so that they could be clearly heard through the recording. That they could be clearly heard was visually indicated to them at the time of the recording by the movement of light on the recording device. No time limit was specified for completion of the recordings.

Instructions: The evaluations will be done on a chart designed on an Excel file that matches the list of items read on the recordings. Each item on the chart is to be evaluated in terms of Consonant Sound (C), Vowel Sound (V), and Tone (T) under the columns marked (C), (V), and (T). On the first Excel Worksheet, the consonant, vowel, and tone of the monosyllabic words will be rated. On the second Worksheet, the same consonant, vowel, and tone highlighted in red will be rated as they occur in disyllabic expressions. There are ten types of target syllables to be evaluated, namely, ji, qi, xi, zhi, chi, shi, ri, and zi, ci, si. All syllable-types other than these ten types are non-target syllables.

***Important note about non-target syllables in disyllabic expressions:** Please ignore pronunciation errors made on non-target syllables. In other words, pronunciation errors made in non-target syllables should not influence the rating speakers receive for the target syllables. For example, if the disyllabic expression is 'hai zi', but the subject reads 'hei zi', the rating for that expression must only reflect the degree of similarity or lack of similarity of the subject's pronunciation to 'zi'. The mistake in 'hai' is ignored, because it is a non-target syllable, falling outside the scope of this study.

Each consonant (C) and vowel (V) sound will be rated on a three-point scale based on its similarity to the target consonant or vowel sound of Standard Mandarin Chinese, known as Putonghua. 1 represents the lowest possible rating and 3 represents the highest possible rating.

1= dissimilar to target sound; unrecognizable to a native speaker likely to cause a failure to communicate

2= somewhat similar to target sound; still acceptable to a native speaker for communication, although possibly distracting to communication

3= very similar to target sound; easily understood by and acceptable to a native speaker

A dash '–', to represent 'no rating was given for an item', can be supplied for an item that has been skipped over by the subject.

Each tone (T) will be rated on a three-point scale based on its similarity to target-likeness as follows.

1= dissimilar to target sound; unrecognizable to a native speaker likely to cause a failure to communicate

2= somewhat similar to target sound; still acceptable to a native speaker for communication, although possibly distracting to communication

3= very similar to target sound; easily understood by and acceptable to a native speaker

APPENDIX E. Results of CFL Learner Survey on Attitudes toward Pronunciation

True Beginner Monolingual Group: Ave. Age 15.3 yrs / 6 Male, 0 Female

RATING							
Total	1	2	3	4	5	#	Item
6	0	2	2	2	0	*1.*	Pronunciation can be easily acquired.
6	0	0	0	3	3	*2.*	I consider learning a native-like pronunciation an important goal of mine.
6	0	0	2	3	1	*3.*	I feel comfortable speaking the target language in front of my peers.
6	0	1	3	2	0	*4.*	Learning pronunciation is fun.
6	2	1	0	2	1	*5.*	Pronunciation is determined by birth, origin, and mother tongue.
6	0	0	1	5	0	*6.*	Pronunciation enhances social relationships.
36	**2**	**4**	**8**	**17**	**5**		

Non-beginner Monolingual Group: Ave. Age 16.8 yrs / 4 Male, 1 Female

RATING							
Total	1	2	3	4	5	#	Item
5	0	2	1	2	0	*1.*	Pronunciation can be easily acquired.
5	1	0	0	3	1	*2.*	I consider learning a native-like pronunciation an important goal of mine.
5	0	1	0	3	1	*3.*	I feel comfortable speaking the target language in front of my peers.
5	0	2	1	2	0	*4.*	Learning pronunciation is fun.
5	0	2	2	0	1	*5.*	Pronunciation is determined by birth, origin, and mother tongue.
5	0	0	1	3	1	*6.*	Pronunciation enhances social relationships.
30	**1**	**7**	**5**	**13**	**4**		

True Beginner Asian Bilingual Group: Ave. Age 16 yrs / 3 Male, 2 Female

RATING							
Total	1	2	3	4	5	#	Item
5	0	1	3	0	1	*1.*	Pronunciation can be easily acquired.
5	0	0	2	2	1	*2.*	I consider learning a native-like pronunciation an important goal of mine.
5	0	3	0	2	0	*3.*	I feel comfortable speaking the target language in front of my peers.
5	1	0	4	0	0	*4.*	Learning pronunciation is fun.
5	1	0	0	0	4	*5.*	Pronunciation is determined by birth, origin, and mother tongue.
5	0	0	1	1	3	*6.*	Pronunciation enhances social relationships.
30	**2**	**4**	**10**	**5**	**9**		

Heritage Group: Ave. Age 16.3 yrs / 5 Male, 9 Female

RATING							
Total	1	2	3	4	5	#	Item
14	3	4	4	2	1	*1.*	Pronunciation can be easily acquired.
14	0	0	2	4	8	*2.*	I consider learning a native-like pronunciation an important goal of mine.
14	0	1	3	6	4	*3.*	I feel comfortable speaking the target language in front of my peers.
14	0	1	3	3	7	*4.*	Learning pronunciation is fun.
14	2	1	0	7	4	*5.*	Pronunciation is determined by birth, origin, and mother tongue.
14	0	2	4	6	2	*6.*	Pronunciation enhances social relationships.
84	**5**	**9**	**16**	**28**	**26**		

All Groups: Ave. Age. 16.2 yrs / 18 Male, 12 Female

RATING							
Total	1	2	3	4	5	#	Item
30	3	9	10	6	2	*1.*	Pronunciation can be easily acquired.
30	1	0	4	12	13	*2.*	I consider learning a native-like pronunciation an important goal of mine.
30	0	5	5	14	6	*3.*	I feel comfortable speaking the target language in front of my peers.
30	1	4	11	7	7	*4.*	Learning pronunciation is fun.
30	5	4	2	9	10	*5.*	Pronunciation is determined by birth, origin, and mother tongue.
30	0	2	7	15	6	*6.*	Pronunciation enhances social relationships.
180	**10**	**24**	**39**	**63**	**44**		

BM : Beginner Monolingu NM: Non-beginner Monolingual BA: Beginner Asian Bilingual H: Heritage

APPENDIX F. Comparative Chart of Average Scores by Target Sound and Learner Group

#	Pinyin	Pinyin	W-BM	P-BM	Aver.-BM	W-H	P-H	Aver.-H	W-NM	P-NM	Aver.-NM	W-BA	P-BA	Aver.-BA	Aver.-4 Groups	Aver.-4 G / One Sound (C+V+T)	Aver.-4 G / One Sound (C+V)
A1	zĪ	hái zǐ/zi	C	C	C	C	C	C	C	C	C	C	C	C	C		
A2	zǐ	hán zi	92.6	98.1	95.4	99.2	98.4	98.8	86.7	80.0	83.3	84.4	81.1	82.8	90.1		
A3	zĪ	gōng zī	V	V	V	V	V	V	V	V	V	V	V	V	V	zi	zi
A4	zĪ	zǐ sè	82.4	84.3	83.3	92.5	92.9	92.7	100.0	95.6	97.8	77.8	72.2	75.0	87.2	86.4	88.7
A5	zǐ	zì jǐ	T	T	T	T	T	T	T	T	T	T	T	T	T		
A6	zĪ	zī shì	80.6	71.3	75.9	96.0	88.5	92.3	97.8	92.2	95.0	58.9	70.0	64.4	81.9		
B1	zhĪ	zhī piào	C	C	C	C	C	C	C	C	C	C	C	C	C		
B2	zhĪ	yù zhī	88.9	89.8	89.4	88.9	80.2	84.5	100.0	98.9	99.4	73.3	78.9	76.1	87.4		
B3	zhí	zhí zǒu	V	V	V	V	V	V	V	V	V	V	V	V	V	zhi	zhi
B4	zhĪ	bù zhī	88.9	90.7	89.8	92.1	84.5	88.3	100.0	98.9	99.4	94.4	92.2	93.3	92.7	87.3	90.1
B5	zhĪ	bái zhǐ	T	T	T	T	T	T	T	T	T	T	T	T	T		
B6	zhĪ	zhī dìng	87.0	66.7	76.9	90.1	82.1	86.1	97.8	91.1	94.4	65.6	74.4	70.0	81.9		
C1	jĪ	jī gè/ge	C	C	C	C	C	C	C	C	C	C	C	C	C		
C2	jĪ	fēi jī	94.4	84.3	89.4	100.0	99.2	99.6	100.0	100.0	100.0	92.2	86.7	89.4	94.6		
C3	jĪ	jī ròu	V	V	V	V	V	V	V	V	V	V	V	V	V	ji	ji
C4	jĪ	wàng jì	88.0	88.0	88.0	100.0	99.2	99.6	100.0	100.0	100.0	81.1	75.6	78.3	91.5	88.5	93.1
C5	jĪ	jì shì	T	T	T	T	T	T	T	T	T	T	T	T	T		
C6	jĪ	nián jì	86.1	68.5	77.3	91.7	87.3	89.5	93.3	82.2	87.8	57.8	68.9	63.3	79.5		
A1.1	cí	shēng cí	C	C	C	C	C	C	C	C	C	C	C	C	C		
A2.1	cī	sān cì	56.5	62.0	59.3	82.5	82.9	82.7	86.7	81.1	83.9	42.2	44.4	43.3	67.3		
A3.1	cí	cí qì	V	V	V	V	V	V	V	V	V	V	V	V	V	ci	ci
A4.1	cì	cì' ěr	64.8	71.3	68.1	74.2	85.7	80.0	100.0	88.9	94.4	53.3	62.2	57.8	75.1	71.8	71.2
A5.1	cí	cí tiě	T	T	T	T	T	T	T	T	T	T	T	T	T		
A6.1	cí	gào cí	73.1	63.9	68.5	84.5	81.3	82.9	95.6	81.1	88.3	52.2	52.2	52.2	73.0		
B1.1	chĪ	chī fàn	C	C	C	C	C	C	C	C	C	C	C	C	C		
B2.1	chĪ	yá chǐ	87.0	98.1	92.6	96.8	94.8	95.8	97.8	92.2	95.0	95.6	100.0	97.8	95.3		
B3.1	chí	chí táng	V	V	V	V	V	V	V	V	V	V	V	V	V	chi	chi
B4.1	chĪ	yīng chī	90.7	99.1	94.9	94.0	95.6	94.8	97.8	95.6	96.7	88.9	100.0	94.4	95.2	90.1	95.3
B5.1	chī	tāng chī	T	T	T	T	T	T	T	T	T	T	T	T	T		
B6.1	chī	chī jiāo	85.2	69.4	77.3	87.3	89.3	88.3	88.9	93.3	91.1	55.6	68.9	62.2	79.7		
C1.1	qĪ	shí qī	C	C	C	C	C	C	C	C	C	C	C	C	C		
C2.1	qí	qí guài	99.1	90.7	94.9	98.4	91.3	94.8	100.0	100.0	100.0	100.0	95.6	97.8	96.9		
C3.1	qí	shēng qí	V	V	V	V	V	V	V	V	V	V	V	V	V	qi	qi
C4.1	qí	qí shí	100.0	96.3	98.1	98.4	96.4	97.4	100.0	100.0	100.0	93.3	91.1	92.2	96.9	91.0	96.9
C5.1	qí	xīng qī	T	T	T	T	T	T	T	T	T	T	T	T	T		
C6.1	qī	qī chuáng	75.0	69.4	72.2	88.9	84.9	86.9	86.7	93.3	90.0	65.6	70.0	67.8	79.2		
A1.2	sī	sì shí	C	C	C	C	C	C	C	C	C	C	C	C	C		
A2.2	sī	shēng sī	96.3	83.3	89.8	98.8	87.3	93.1	97.8	95.6	96.7	70.0	92.2	81.1	90.2		
A3.2	sī	yì sī/si	V	V	V	V	V	V	V	V	V	V	V	V	V	si	si
A4.2	sī	gōng sī	70.4	77.8	74.1	84.5	82.9	83.7	100.0	97.8	98.9	63.3	81.1	72.2	82.2	85.3	86.2
A5.2	sī	sī chóu	T	T	T	T	T	T	T	T	T	T	T	T	T		
A6.2	sī	sī wén	88.9	70.4	79.6	97.6	84.9	91.3	100.0	98.9	99.4	68.9	58.9	63.9	83.6		
B1.2	shí	shí wàn	C	C	C	C	C	C	C	C	C	C	C	C	C		
B2.2	shĪ	lǎo shī	100.0	98.1	99.1	98.0	99.2	98.6	100.0	97.8	98.9	86.7	94.4	90.6	96.8		
B3.2	shĪ	kāi shǐ	V	V	V	V	V	V	V	V	V	V	V	V	V	shi	shi
B4.2	shí	shí qing	100.0	98.1	99.1	98.0	99.6	98.8	100.0	97.8	98.9	86.7	94.4	90.6	96.8	91.6	96.8
B5.2	shì	jiào shì	T	T	T	T	T	T	T	T	T	T	T	T	T		
B6.2	shì	shì zi	85.2	71.3	78.2	92.1	89.3	90.7	95.6	88.9	92.2	60.0	66.7	63.3	81.1		
C1.2	xĪ	dōng xī	C	C	C	C	C	C	C	C	C	C	C	C	C		
C2.2	xĪ	xī shōu	74.1	94.4	84.3	98.8	99.2	99.0	100.0	100.0	100.0	68.9	98.9	83.9	91.8		
C3.2	xĪ	xī huān	V	V	V	V	V	V	V	V	V	V	V	V	V	xi	xi
C4.2	xĪ	xī wàng	99.1	97.2	98.1	99.2	99.2	99.2	100.0	100.0	100.0	66.7	100.0	83.3	95.2	90.4	93.5
C5.2	xí	fù xí	T	T	T	T	T	T	T	T	T	T	T	T	T		
C6.2	xì	yóu xì	87.0	72.2	79.6	91.3	92.5	91.9	94.4	94.4	94.4	68.9	72.2	70.6	84.1		
B1.3	rì	rì lì	C	C	C	C	C	C	C	C	C	C	C	C	C		
		jié rì	94.4	86.1	90.3	81.0	78.6	79.8	86.7	85.6	86.1	60.0	61.1	60.6	79.2		
		rì wén	V	V	V	V	V	V	V	V	V	V	V	V	V	ri	ri
		rì jì	100.0	100.0	100.0	100.0	97.2	98.6	93.3	100.0	96.7	73.3	74.4	73.9	92.3	86.6	85.7
		míng rì	T	T	T	T	T	T	T	T	T	T	T	T	T		
		shēng rì	94.4	73.1	83.8	88.1	99.6	93.8	100.0	100.0	100.0	80.0	72.2	76.1	88.4		
	Total Aver.		87.0	82.8	84.9	92.8	90.8	91.8	96.6	94.0	95.3	72.9	78.4	75.6	86.9	86.9	89.7

	Aver.	W-BM	P-BM	Aver.-BM	W-H	P-H	Aver.-H	W-NM	P-NM	Aver.-NM	W-BA	P-BA	Aver.-BA	Aver.-4 Groups
All Ten Sounds	10 C	88.3	88.5	88.4	94.2	91.1	92.7	95.6	93.1	94.3	77.3	83.3	80.3	88.9
	3 V	88.4	90.3	89.4	93.3	93.3	93.3	99.1	97.4	98.3	77.9	84.3	81.1	90.5
	4 T	84.3	69.6	76.9	90.8	88.0	89.4	95.0	91.6	93.3	63.3	67.4	65.4	81.2

W-BM (Words-Beginner Monolingual) *W- H (Words-Heritage)* *W-NM (Words-Nonbeginner Monolingual)* *W-BA (Words-Beginner Asian)*
P- BM (Phrases- Beginner Monolingual) *P-H (Phrases-Heritage)* *P-NM (Phrases-Nonbeginner Monolingual)* *P-BA (Phrases-Beginner Asian)*

C = Consonant *V = Vowel* *T = Tone*

APPENDIX G. Learners' Perceptions of Sound Difficulty

Syllables that learners identified as most difficult

	Dentals			Palatals				Retroflexes		
Broups	**zi**	**ci**	**si**	**ji**	**qi**	**xi**	**zhi**	**chi**	**shi**	**ri**
BM	0	3	1	0	1	2	1	0	0	2
NM	0	1	1	0	0	0	0	0	0	0
BA	2	3	1	0	0	1	0	0	0	2
H	3	4	3	1	0	0	2	0	0	6
TOTAL	**5**	**11**	**6**	**1**	**1**	**3**	**3**	**0**	**0**	**10**

Syllable-types that learners ranked as most difficult

Groups	Dentals	Palatals	Retroflexes
BM	2	2	2
NM	3	2	0
BA	4	1	0
H	7	3	4
TOTAL	**16**	**8**	**6**

Syllable-types ranked as most difficult by learners that correctly categorized all syllables by vowels

Dentals	Palatals	Retroflexes
0	1	1
2	1	0
1	0	0
4	2	3
7	**4**	**4**

Syllble-types that learners ranked as least difficult

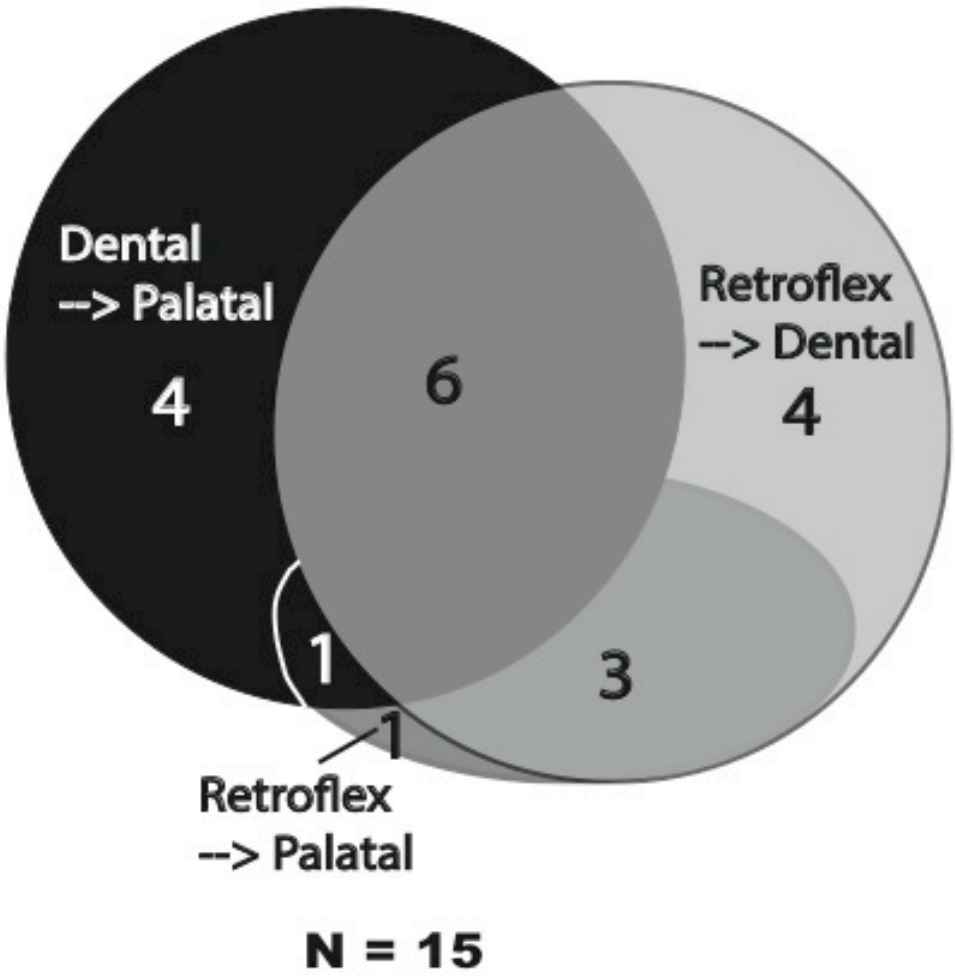
Groups	Dentals	Palatals	Retroflexes
BM	0	3	3
NM	1	3	1
BA	0	2	3
H	3	6	5
TOTAL	**4**	**14**	**12**

Syllable-types ranked as least difficult by learners that correctly categorized all syllables by vowels

Dentals	Palatals	Retroflexes
0	1	1
1	2	0
0	0	1
1	4	4
2	**7**	**6**

'M : Beginner Monolingual NM: Non-beginner Monolingual BA: Beginner Asian Bilingual H: Heritage

APPENDIX H. Syllable-types Conflated by Learners on the Syllable-grouping Task



Notes:
(1) '-->' stands for 'was conflated with' where X --> Y.
(2) 'Dental' refers to dental syllables: 'zi', 'ci', 'si'.
(3) 'Retroflex' refers to retroflex syllables: 'zhi', 'chi', 'shi', 'ri'.
(4) 'Palatal' refers to palatal syllables: 'ji', 'qi', 'xi'.
(5) The Syllable-grouping task prompted learners to group ten target syllables into three groups based on likeness of vowel sound.

APPENDIX I. Results of Diagramming Task

Articulators involved	ci	chi	qi
Tip of tongue	4	0	0
Blade of tongue	4	7	7
Root of tongue	0	0	0
Nose	0	0	0
Lips	3	2	5
Upper teeth	7	2	2
Bottom teeth	5	5	3
Alveolar ridge	9	6	6
Hard palate	1	7	8
Soft palate	1	1	1
Pharynx	1	2	1
Uninterpretable	3	3	3
TOTAL	38	35	36

(N=30) *Some learners drew more than one arrow for each sound.

Discrimination of Articulatory Locus of Sounds	Student #
Front-to-back relation on roof of mouth	10
Front-to-back relation on tongue	1
No front-to-back relation recognized	19

Diagram of the Articulators Used in Speech

BM : Beginner Monolingual *NM: Non-beginner Monolingual*
BA: Beginner Asian Bilingual *H: Heritage*

APPENDIX J. Types of L1 Sounds Transferred for 'ri' on L1 Sound-transfer Task

L1 sound / Group	No L1 equivalent	Affricates	Fricatives	Initial or final English r	Other	TOTAL
BM	3	0	3	0	0	6
BA	0	2	1	1	1	5
NM	1	2	1	1	0	5
H	7	0	2	5	0	14
Total	11	4	7	7	1	30
Percentage	37%	13%	23%	23%	3%	100%
Example words of each type of sound	*n/a*	*puncture, danger*	*measure, shirt, sir*	*rent, her*	*drum*	--

APPENDIX K. Strategies Used by Learners in Distinguishing 'zi' and 'zhi'

Type of Explanation Given by the Learners	BM	NM	BA	H	TOTAL	Examples
(1) Not sure/No explanation	1	0	1	2	4	Not sure the diffrence :)'
(2) Comparison to English sounds involved	3	4	4	5	16	zi: z as in zebra i: I as in ignite'
(3) Reference to articulators involved	0	0	0	3	3	'zi uses tip of tongue. Zhi is where the tongue touches top of mouth.
(4) Manner of articulation involved	1	1	0	2	4	'For zi your mouth is like a straight line while the zhi is like a circle. And for zi your tongue is basically flat while the zhi your tongue is more relaxed and rounded.
(5) Description of quality of sound	1	0	0	1	2	'The z has more buzz to it, while the zh sounds flat.'
(6) Reference to tones involved	0	0	0	1	1	'zi is fourth tone. Zhi is second tone.
TOTAL	6	5	5	14	30	